Finite Fields and Their Applications 32 (2015) 120–147

Contents lists available at ScienceDirect

Finite Fields and Their Applications

www.elsevier.com/locate/ffa

Perfect nonlinear functions and cryptography

Céline Blondeau, Kaisa Nyberg

Aalto University, School of Science, Department of Information and Computer Science, Finland

ARTICLE INFO

Article history:
Received 20 March 2014
Received in revised form 6 October 2014
Accepted 10 October 2014
Available online 7 November 2014
Communicated by Gary McGuire

MSC:
11T71
94A60

Keywords:
Perfect nonlinear functions
PN functions
Almost perfect nonlinear functions
APN functions
Differential uniformity
Nonlinearity
Differential cryptanalysis

ABSTRACT

In the late 1980s the importance of highly nonlinear functions in cryptography was first discovered by Meier and Staffelbach from the point of view of correlation attacks on stream ciphers, and later by Nyberg in the early 1990s after the introduction of the differential cryptanalysis method. Perfect nonlinear (PN) and almost perfect nonlinear (APN) functions, which have the optimal properties for offering resistance against differential cryptanalysis, have since then been an object of intensive study by many mathematicians. In this paper, we survey some of the theoretical results obtained on these functions in the last 25 years. We recall how the links with other mathematical concepts have accelerated the search on PN and APN functions. To illustrate the use of PN and APN functions in practice, we discuss examples of ciphers and their resistance to differential attacks. In particular, we recall that in cryptographic applications suboptimal functions are often used.

1. Introduction

The derivative of a real or complex valued function is a useful tool when studying various mathematical and physical phenomena. By definition, the derivative of

E-mail addresses: celine.blondeau@aalto.fi (C. Blondeau), kaisa.nyberg@aalto.fi (K. Nyberg).

http://dx.doi.org/10.1016/j.ffa.2014.10.007

a differentiable function at a given point provides the best affine approximation of the function. For functions defined over finite groups the notion of derivative takes a different appearance and is closely related to designs and combinatorial structures such as, for example, difference sets [1]. If the domain of definition of the function is a linear space over a finite field, then also in this case a close connection between derivatives of the function and its linear approximations can be established as we will see later in this paper.

In the late 1980s, new approaches to the cryptanalysis of block ciphers were introduced. In his study of FEAL-4, Sean Murphy [2] exploited solutions of equations of the form $G(x+a)+G(x+b)=d$. About at the same time, Eli Biham and Adi Shamir [3] studied the block cipher DES and showed that for some fixed plaintext differences, certain differences in the encrypted values appear much more often than one would expect on average. Furthermore, they showed how one can exploit this phenomenon to recover information on the secret key. These attacks have launched a lot of interest in the derivatives of functions defined over finite spaces with the goal to mitigate the threat of differential cryptanalysis. While in the design of practical ciphers it is not necessary (and is sometimes even harmful) that the values of the derivatives are optimally distributed, also the functions with optimal derivatives, known as perfect nonlinear or almost perfect nonlinear, have drawn a lot of attention. The discovery in 2009 of an APN permutation in a field of characteristic 2 and even dimension [4] has brought new motivation and new ideas to this field of research.

The selection of results on PN and APN functions presented in this paper is not exhaustive. In particular, we would like to apologize if some important results are missing. Other surveys on APN functions can be found in for instance [5,6].

The rest of the paper is organized as follows. We start in Section 2 by introducing the basic definitions. In Section 3, we introduce some further notions such as bentness that are closely linked with the notions of perfect nonlinear (PN) and almost perfect nonlinear (APN) functions. The link with linear codes is also briefly summarized. Section 3.3 is dedicated to the classes of equivalence which preserve the differential properties. In Section 4, some classical results on PN and APN monomial and polynomial functions are summarized. In particular, the relation between the only known APN permutation over $\mathbb{Z}_2^6$ and quadratic APN polynomials is recalled. Section 5 is dedicated to the exponential and logarithmic functions and on the recent results on the linearity of related functions. In Section 6 we discuss several ciphers, and the use of PN or APN functions in practice. Different approaches to the design and cryptanalysis are considered in this section. Section 7 concludes this paper.

2. Preliminaries

In this paper, we denote by A or B an Abelian group and by $\mathbb{Z}_q^n$ a Cartesian product of n copies of the ring $\mathbb{Z}_q$, where q is a positive integer greater than 1. The results in

the case of $q = 2$ may take essentially different form than the ones in the case of other values of q and hence a separate treatment is often required.

When working with functions of several variables in $\mathbb{Z}_q$, we will denote by f a q-ary function with range in $\mathbb{Z}_q$. Then $f : \mathbb{Z}_q^n \to \mathbb{Z}_q$, and if $q = 2$ the function f is a Boolean function. When the range of the function is $\mathbb{Z}_q^m$, where $m > 1$, we will use a capital letter F to denote $F : \mathbb{Z}_q^n \to \mathbb{Z}_q^m$. In the case of $q = 2$, such a function is called a vectorial Boolean function. Capital letters will also be used to denote functions defined in general Abelian groups.

In a finite domain, there is no question about differentiability of a function as the derivative always exists and is defined as follows.

Definition 1. Let A and B be finite Abelian groups and $F : A \to B$ be a function. Given $a \in A$ the function defined by

$$\begin{aligned} D_a F : \quad & A \to B \\ & x \mapsto F(x+a) - F(x) \end{aligned}$$

is called a derivative of F. Given $a \in A$ and $b \in B$ the relation

$$F(x+a) - F(x) = b \tag{1}$$

is called a differential of F with input difference a and output difference b.

Already since the end of the 1960s, derivatives of functions $F : A \to B$ such that $|A| = |B|$ were studied in [7]. In particular, functions with bijective derivatives, called as *planar* functions, received attention. The introduction of differential cryptanalysis served as a motivation to study differentials of nonlinear functions and upperbounds to the number of solutions to Eq. (1).

Definition 2. (See [8].) Let $F : A \to B$ be a function and set

$$\delta(a,b) = \left|\left\{x \mid F(x+a) - F(x) = b\right\}\right|.$$

We denote by Δ_F the positive integer defined as

$$\Delta_F = \max_{\substack{a \in A, a \neq 0 \\ b \in B}} \delta(a,b).$$

Then F is said to be differentially Δ_F-uniform.

Clearly, we always have $\Delta_F \geq |B|/|A|$. We say that F is *perfect nonlinear* (PN), if $\Delta_F = |B|/|A|$. This notion was studied for Boolean functions by Willi Meier and Othmar Staffelbach [9], who coined the term perfect nonlinearity for Boolean functions

of n variables satisfying $\Delta_f = 2^{n-1}$. It was soon discovered that the perfect nonlinearity is equivalent to the bentness introduced already in the 1970s by Oscar Rothaus [10], an NSA mathematician and Cornell mathematics professor, who is more widely known for his contributions to the development of the Hidden Markov Model.

The studies of nonlinearity of Boolean functions by Meier and Staffelbach were motivated by the cryptanalysis of stream ciphers. The introduction of differential cryptanalysis of the block cipher DES by Biham and Shamir [3] raised the need to study this concept for S-boxes and nonlinear round functions which are nothing else than vectorial Boolean functions [11]. The following result about the value distribution of a perfect nonlinear function was proved in [11].

Theorem 1. *(See [11].) Let $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$ be PN. Then, for each y in the image set of F,*

$$\left|F^{-1}(y)\right| = \left|\left\{x \in \mathbb{Z}_2^n \mid F(x) = y\right\}\right| = b_y 2^{\frac{n}{2}-m},$$

where b_y is an odd integer.

As $|F^{-1}(y)|$ is a non-zero integer for at least some $y \in \mathbb{Z}_2^m$, it follows that PN functions from $\mathbb{Z}_2^n$ to $\mathbb{Z}_2^m$ exist only if $n \geq 2m$.

The definition of bent functions was extended to q-ary functions in [12], and for any prime q, a function is perfect nonlinear if and only if it is generalized bent. Theorem 1 also extends itself to the general q-ary case [11].

In the case of general Abelian groups, where $|A| = |B|$, a perfect nonlinear function is also called *planar*. As the derivatives of a planar function are bijective, there exists one $x \in A$ such that $F(x+a)-F(x) = 0$. It follows that a planar function is never one-to-one. On the other hand, for vectorial Boolean functions, if Eq. (1) has one solution, x, it has at least two solutions, $(x, x+a)$. Hence the minimum value for the differential uniformity is 2. If this minimum value is achieved, a function is called *almost perfect non-linear* (APN). More generally, a function $F : A \to B$ is called APN if it is differentially $\frac{2|B|}{|A|}$-uniform.

While probably without direct application in cryptography a new definition of planar function has recently been introduced for vectorial Boolean function. In [13], a function $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is called planar if for all non-zero $a \in \mathbb{Z}_2^n$ the functions $x \mapsto F(x) + F(x+a)+ax$ are bijective. This notion which is not related to the definitions of PN and APN functions studied in this paper will not be further discussed.

3. Nonlinearity and coding theory

While most of the results concerning PN and APN functions can be generalized to any Abelian groups, we focus in this section on the special Abelian group formed by taking the Cartesian product $\mathbb{Z}_q^n$ of n copies of the additive group $\mathbb{Z}_q$. In addition, we often assume that q is a prime, in which case $\mathbb{Z}_q^n$ is a linear space and can be identified with the field $\mathbb{F}_{q^n}$.

3.1. Bent and almost bent functions

Let $q > 1$ be an integer and let us denote by $w \in \mathbb{C}$ the q-th root of unity, that is, $w = e^{2i\pi/q}$. The *Walsh transform* of a q-ary function $f : \mathbb{Z}_q^n \to \mathbb{Z}_q$ computed in $\mathbb{C}$ is defined as

$$\hat{f}(a) = \sum_{x \in \mathbb{Z}_q^n} w^{f(x) - \langle a, x \rangle}$$

where $a \in \mathbb{Z}_q^n$ and the notation $\langle a, x \rangle$ means the sum of the products $a_i x_i$ in $\mathbb{Z}_q$, for $a = (a_1, \ldots, a_n) \in \mathbb{Z}_q^n$ and $x = (x_1, \ldots, x_n) \in \mathbb{Z}_q^n$. If $q = 2$, we have $w = -1$. For any prime $q = p$, the sum $\langle a, x \rangle$ is a scalar product in the linear space $\mathbb{Z}_p^n$, in which case the scalar product $\langle a, x \rangle = \mathbf{Tr}(ax)$ defined by the absolute trace function $y \mapsto \mathbf{Tr}(y) = \sum_{i=0}^{n-1} y^{p^i}$ in $\mathbb{Z}_p$ is a natural choice.

The values of the Walsh transform $\hat{f}(a)$ are called the *Walsh coefficients* of f. They are used to measure the distance of f to the functions $x \mapsto \langle a, x \rangle$, $a \in \mathbb{Z}_p^n$. For $q = p$ prime, these are exactly the linear functions in the linear space $\mathbb{Z}_p^n$. We define the *linearity* $\mathcal{L}(f)$ of f as

$$\mathcal{L}(f) = \max_{a \in \mathbb{Z}_q^n} \left| \hat{f}(a) \right|.$$

The *linearity* of a vectorial function $F : \mathbb{Z}_q^n \to \mathbb{Z}_q^m$ is defined based on the linearities of the non-zero linear combinations of its coordinate functions. Given $\lambda \in \mathbb{Z}_q^m \setminus \{0\}$, the q-ary function $f_\lambda(x) = \langle \lambda, F(x) \rangle$ is called the λ-*component* of F. The *Walsh spectrum* of F is then the set of all values of $\hat{f_\lambda}(a)$ for all $a \in \mathbb{Z}_q^n$ and $\lambda \in \mathbb{Z}_q^m \setminus \{0\}$. The linearity $\mathcal{L}(F)$ of the vectorial function F is then defined as

$$\mathcal{L}(F) = \max_{\lambda \in \mathbb{Z}_q^m \setminus \{0\}} \mathcal{L}(f_\lambda).$$

For any function $F : \mathbb{Z}_q^n \to \mathbb{Z}_q^m$, we have

$$q^{n/2} \leq \mathcal{L}(F) \leq q^n.$$

The upperbound is achieved by the functions $x \mapsto \langle a, x \rangle$. The lower bound follows from Parseval's theorem. A function which achieves the lower bound is called *bent* if $q = 2$ and *generalized q-ary bent* in the general case [12]. When $m = n$, bent functions do not exist, see Theorem 1, which also holds under some additional assumptions for generalized q-ary bent functions.

For odd n and $q = 2$, the lower bound for $\mathcal{L}(F)$ is $2^{(n+1)/2}$ and functions that achieve this bound are known to exist [14]. Such functions are called *almost bent* (AB). For even n, it is possible to obtain a function $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ with $\mathcal{L}(F) = 2^{n/2+1}$ though it is still an open problem whether this value is the minimum possible.

For bijective functions $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ the minimum known value of $\mathcal{L}(F)$ is $2^{\lfloor n/2 \rfloor + 1}$. A list of known power permutations on $\mathbb{F}_{2^n}$ with the best known nonlinearity can be found in [15], see also Section 4.1. Among them is the inversion monomial. More generally, it has been a longstanding conjecture (see [16] for the characteristic 2) that all monomial functions are such that $\mathcal{L}(F) \geq q^{n/2+1}$.

One of the main tools Chabaud and Vaudenay used in [14] for proving useful and nontrivial results on lower bounds of linearity was the following theorem, which links the derivatives and linear approximations of a vectorial Boolean functions. Recently this result was used in [17,18] in the cryptographic context to show relations between statistical distinguishers in the differential and linear cryptanalysis contexts.

Theorem 2. *(See [14].) Let $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$ be a vectorial Boolean function. Let $\delta(a, b)$ be defined as in Definition 2. Then, $\delta(a, b)$ can be computed as*

$$\delta(a,b) = \frac{1}{2^{2n}} \sum_{u \in \mathbb{Z}_2^n} \sum_{v \in \mathbb{Z}_2^m} \sum_{x \in \mathbb{Z}_2^n} \sum_{y \in \mathbb{Z}_2^n} (-1)^{\langle u, a+x+y \rangle + \langle v, b+F(x)+F(y) \rangle}. \tag{2}$$

Already in [10] it was shown that a Boolean function is bent if and only it is PN. For q-ary functions PN implies bent and the converse is true if q is prime [12]. Naturally, the same implications and equivalences hold for vectorial functions. The following result about the relationships between PN and APN properties can be obtained as a direct application of Theorem 2.

Theorem 3. *Let $F : \mathbb{Z}_2^n \mapsto \mathbb{Z}_2^m$. If F is AB then F is APN. If F is APN and all Walsh coefficients of $\langle 1, F(x) \rangle$ are divisible by 2^{m+1}, then F is AB.*

This link between AB and APN functions has been used in the literature to find new APN functions or to determine the differential uniformity or linearity of a function.

Later, the definitions of bent and PN functions presented in this subsection have been generalized to arbitrary Abelian groups [19]. In particular, a definition of bentness, which is equivalent to the PN property, is given in [20]. Let us present here the special case of additive Abelian groups $\mathbb{Z}_q$ and $\mathbb{Z}_r$. Let us denote by $w_q \in \mathbb{C}$ and $w_r \in \mathbb{C}$ the q-th and r-th roots of unity in $\mathbb{C}$. Then $x \mapsto w_q^{ax}$ defines a character in $\mathbb{Z}_q$ for all $a \in \mathbb{Z}$, whereby it suffices to restrict to $0 \leq a < q$. Similarly, we consider the set of characters $y \mapsto w_r^{by}$, $0 \leq b < r$, in $\mathbb{Z}_r$. Then a function $F : \mathbb{Z}_q \to \mathbb{Z}_r$ is called *bent* if for all $0 \leq a < q$ and $0 < b < r$

$$\left| \sum_{x \in \mathbb{Z}_q} w_r^{bF(x)} \overline{w}_q^{ax} \right| = \sqrt{q},$$

where $\overline{w}_q$ is the complex conjugate of w_q. The linearity of a function $F : \mathbb{Z}_q \to \mathbb{Z}_r$ is defined analogically as above as

$$\mathcal{L}(F) = \max_{0 \le a < q, 0 < b < r} \left| \sum_{x \in \mathbb{Z}_q} w_r^{bF(x)} \overline{w}_q^{ax} \right|.$$

Then $\sqrt{q} \le |\mathcal{L}(F)| \le q$.

3.2. Linear codes

After the importance of nonlinearity and APN functions for cryptography was discovered, many results previously found in the coding theory context obtained new interpretations. In this section, we recall how the minimal distance of a certain linear code is linked with PN and APN functions. Let $F : \mathbb{Z}_q^n \to \mathbb{Z}_q^m$ be a q-ary function and let H_F the $(1 + n + m) \times q^n$ matrix defined as follows:

$$H_F = \begin{pmatrix} \cdots & 1 & \cdots \\ \cdots & x & \cdots \\ \cdots & F(x) & \cdots \end{pmatrix}_{x \in \mathbb{Z}_q^n},$$

the code C_F defined over $\mathbb{Z}_q$ of the vectorial function F is the linear code with parity check matrix H_F.

Results on the minimal distance (or weight enumerator) of this code C_F have been derived depending on the property of the derivative of F. In particular, in characteristic 2, a function $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is APN if and only if the code C_F defined over $\mathbb{Z}_2$ has minimal distance 6 [21]. In the same paper, it is shown that when the function is AB, the related code C_F has a particular weight distribution. The proof is derived from the characterization of APN functions.

While this result does not generalize directly to PN functions, it has been proved that the weight distribution of the codes corresponding to all known PN functions is the same [22,23]. Relation between PN functions and other linear and cyclic codes has also been studied for instance in [24,22]. In particular, it has been shown that the minimal distance of some linear codes defined over $\mathbb{Z}_p$, $p > 2$, can be determined by the PN property of the underlying function F.

Among different other applications, this link with coding theory, has been used in 2009 [4], to find the first APN permutation over $\mathbb{Z}_2^n$, for an even n. More details on this function defined over $\mathbb{F}_{2^6}$ are given in Section 4.2.

3.3. CCZ- and EA-equivalence

As shown in [25] the linearity of a function is preserved when composing the function with a bijective affine function on the left or on the right side. The same statement is true for the differential uniformity. Compositions with non-bijective affine functions do not in general preserve linearity or differential uniformity. Motivated by these properties the following equivalences have been introduced.

Definition 3. Two functions $F, G : \mathbb{Z}_q^n \to \mathbb{Z}_q^m$ are affine equivalent if there exist two bijective affine functions $L_1 : \mathbb{Z}_q^n \to \mathbb{Z}_q^n$ and $L_2 : \mathbb{Z}_q^m \to \mathbb{Z}_q^m$ such that $G(x) = L_2(F(L_1(x)))$ for all $x \in \mathbb{Z}_q^n$. Functions F and G are called extended affine equivalent (EA-equivalent) if, in addition to L_1 and L_2 as above, there exists an affine function $L_3 : \mathbb{Z}_q^n \to \mathbb{Z}_q^m$ such that $G(x) = L_2(F(L_1(x))) + L_3(x)$, for all $x \in \mathbb{Z}_q^n$.

The properties of linearity and differential uniformity are preserved by the EA-equivalence [8,21]. The bijectivity of a function is preserved by the affine equivalence if the affine functions L_1 and L_2 are bijective. The bijectivity of a function is not necessarily preserved by the EA-equivalence.

In [25] and [8] it was also noted that the nonlinearity and differential uniformity of a bijective function and its inverse are equal. With usually different algebraic degrees, these functions are generally not EA-equivalent. In 1998, Claude Carlet, Pascale Charpin, and Victor Zinoviev further generalized the notion of affine equivalence to a property that holds for a function and its inverse.

Definition 4. (See [21,26].) Two functions $F, G : \mathbb{Z}_q^n \to \mathbb{Z}_q^m$ are called CCZ-equivalent if their graphs $\mathcal{G}_F = \{(x, y) \in \mathbb{Z}_q^n \times \mathbb{Z}_q^m \mid y = F(x)\}$ and $\mathcal{G}_G = \{(x, y) \in \mathbb{Z}_q^n \times \mathbb{Z}_q^m \mid y = G(x)\}$ can be transformed from one to another using an affine bijective mapping L in $\mathbb{Z}_q^n \times \mathbb{Z}_q^m$.

In the context of coding theory CCZ-equivalence can be rewritten as follows: two functions F and G are CCZ-equivalent if and only if the codes C_F and C_G are equivalent.

In general, establishing CCZ-equivalence of arbitrary functions is extremely difficult. On the other hand, as many properties are invariant in CCZ-equivalence, it is often possible to determine cases of inequivalence. Functions for which CCZ-equivalence coincides with EA-equivalence are, for instance, planar functions, Boolean functions, vectorial bent functions if $q = p = 2$ [27] and vectorial bent functions if $q = p$ is an odd prime and $n = m$ [28].

4. Polynomials over finite fields

Any function $F : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ can be defined as a polynomial over $\mathbb{F}_{p^n}$ as

$$F(x) = \sum_{i=0}^{p^n-1} a_i x^i,$$

where $a_i \in \mathbb{F}_{p^n}$. In practice, for cryptographic applications, bijective nonlinear functions are widely used. Among all polynomials, monomials, i.e., polynomials with only one non-zero coefficient a_i, have been extensively studied and have been used as the main nonlinear component in cryptographic primitives such as in the AES block cipher [29].

Table 1
Known APN monomials in characteristic 2: $F_d : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, $x \mapsto x^d$.

d	Conditions	Ref.
$2^t + 1$	$1 \leq t \leq n/2$, $\gcd(t, n) = 1$	[30,8]
$2^{2t} - 2^t + 1$	$2 \leq t \leq n/2$, $\gcd(t, n) = 1$	[31]
$2^m + 3$	$n = 2m + 1$	[32,33]
$2^m + 2^{\frac{m}{2}} - 1$	$n = 2m + 1$ and m even	[34,35]
$2^m + 2^{\frac{3m+1}{2}} - 1$	$n = 2m + 1$ and m odd	[34,35]
$2^{n-1} - 1$	n odd	[8,36]
$2^{4g} + 2^{3g} + 2^{2g} + 2^g - 1$	$n = 5g$	[37]

4.1. Monomials

In this section, we denote a monomial as a function $F_d : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$, $x \mapsto x^d$. Monomials, or power functions, are bijective if and only if the exponent d is co-prime with the order $p^n - 1$ of the field. As the mapping $w \mapsto w^{p^i}$ is a linear isomorphism in $\mathbb{F}_{p^n}$, all monomials $F_e(x) = x^e$ with e in the cyclotomic class $\{p^i \cdot d \mid 0 \leq i < n\}$, have the same differential uniformity and nonlinearity as the monomial $F_d(x) = x^d$. Results on monomials are often formulated with regard to the least exponent in the cyclotomic class.

Among all polynomials, monomial functions are mathematically easy to study. In particular, all derivatives have similar properties and the differential uniformity can be computed by studying only one derivative, usually $D_1 F_d$. By a substitution, one can show that for $a \neq 0$, $\delta(a, b) = \delta(1, b/a^d)$. For $b = 0$, the value $\delta(1, 0)$, can be computed as $\delta(1, 0) = \gcd(d, p^n - 1) - 1$, and in particular a monomial function is bijective if and only if $\delta(1, 0) = 0$.

In Table 1, we summarize the list of monomials which have been proved APN in characteristic 2. Not all of these functions are bijective. In particular, it is well known that all APN monomials are bijective for odd n and non-bijective for even n. As the monomials with exponents of Hamming weight 2 are quadratic, their derivatives are linear polynomials, and hence among these functions finding the APN ones is easy. For the inverse monomial, the study of the derivative leads to solving an equation of degree two, hence giving either two or no solutions for n odd, in which case the function is APN. For even n, two more solutions can exist, and the function is differentially 4-uniform [8].

For characteristic $p > 2$ and for $n > 4$, known PN monomials are summarized in Table 2. Similarity between the APN monomials in characteristic 2 and the PN ones in larger characteristic deserves to be noticed. In particular, the derivatives of quadratic functions are 2-to-1 in characteristic 2 and bijective in characteristic $p > 2$.

The question, if the list of PN or APN monomials provided in Tables 1 and 2 is complete, remains an open problem. Research towards answering this question has lead to focusing the study on the exceptional monomials.

Table 2
Known PN monomials in characteristic p odd $F_d : \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$, $x \mapsto x^d$, PN for many n.

d	Conditions	Ref.
2	–	
$p^t + 1$	$n/\gcd(n, t)$ is odd	
$(p^t + 1)/2$	$p = 3$, t is odd, $\gcd(n, t) = 1$	[38]

Definition 5. A PN (resp. APN) monomial F_d defined in a field of characteristic p (resp. 2) is called exceptional if F_d is PN (resp. APN) in $\mathbb{F}_{p^n}$ (resp. $\mathbb{F}_{2^n}$) for infinitely many n. The (positive) exponent d of an exceptional monomial F_d is called an exceptional exponent.

For instance, based on this definition, the inverse function $x \mapsto x^{2^n-2}$ is not exceptional since the value of its positive exponent varies with the size of the field.

Most of the results provided in this area are obtained by studying the number of points of some geometrical structure. This approach of finite geometry was used by Fernando Hernando and Gary McGuire [39] to prove that any infinite class of APN monomials must have a quadratic or a Kasami exponent.

No equivalent of this result in characteristics p has been proved yet, but it is generally believed, as conjectured in [40], that in characteristic p, p odd prime, the only exceptional exponents are 2 and $p^t + 1$, and in addition $(3^t + 1)/2$ in characteristic 3, that is, the exponents given in Table 2. The notion of exceptional planar function can be generalized to any polynomial [41,42]. The most recent classification can be found in [43].

4.2. Polynomial

Finding PN or APN polynomials which are not CCZ-equivalent to a known monomial has been a challenging task in the last few years. For instance, in characteristic 2, the first APN function that is not CCZ-equivalent to any known function was discovered by Yves Edel, Gohar Khureghyan and Alexander Pott in 2005 [44]. This function is defined in the field $\mathbb{F}_{2^{10}}$ as $F(x) = x^3 + \beta \cdot x^{36}$ with $\beta \in \alpha\mathbb{F}_{2^5} \setminus \{0\} \cup \alpha^2\mathbb{F}_{2^5} \setminus \{0\}$ and $\alpha \in \mathbb{F}_{2^{10}}$ of order 3. This discovery has inspired many other works in that direction. For instance, Lilya Budaghyan, Claude Carlet and Gregor Leander [45] have shown that the function $x \mapsto x^3 + \mathbf{Tr}(x^9)$ defined over $\mathbb{F}_{2^n}$ is APN. They also partially demonstrate the CCZ-inequivalence of this function with the other known APN permutations. Later the study of functions of the form $x^d + g(x)$ where x^d is APN have been extended to other exponents d such as $d = 2^n - 2$, the exponent of the inverse function, to study functions of the form $x \mapsto x^{-1} + g(x)$. Most of the currently known PN or APN polynomials are binomial or involve quadratic polynomials. Some examples of such polynomials are provided in Table 3.

For even n, several results indicate that the existence of APN permutations is more limited than for odd n. Already in [46], it is proved that a quadratic permutation of $\mathbb{Z}_2^n$ with n even cannot be APN. Later, Xiang-dong Hou [47] proved the following result.

Table 3
The Banff list of APN polynomials in dimension 6. The element α is primitive in $\mathbb{F}_{2^6}$. These polynomials are not permutation.

1	x^3
2	$x^3 + \alpha^{11}x^6 + \alpha x^9$
3	$\alpha x^5 + x^9 + \alpha^4 x^{17} + \alpha x^{18} + \alpha^4 x^{20} + \alpha x^{24} + \alpha^4 x^{34} + \alpha x^{40}$
4	$\alpha^7 x^3 + x^5 + \alpha^3 x^9 + \alpha^4 x^{10} + x^{17} + \alpha^6 x^{18}$
5	$x^3 + \alpha x^{24} + x^{10}$
6	$x^3 + \alpha^{17}(x^{17} + x^{18} + x^{20} + x^{24})$
7	$x^3 + \alpha^{11}x^5 + \alpha^{13}x^9 + x^{17} + \alpha^{11}x^{33} + x^{48}$
8	$\alpha^{25}x^5 + x^9 + \alpha^{38}x^{12} + \alpha^{25}x^{18} + \alpha^{25}x^{36}$
9	$\alpha^{40}x^5 + \alpha^{10}x^6 + \alpha^{62}x^{20} + \alpha^{35}x^{33} + \alpha^{15}x^{34} + \alpha^{29}x^{48}$
10	$\alpha^{34}x^6 + \alpha^{52}x^9 + \alpha^{48}x^{12} + \alpha^6 x^{20} + \alpha^9 x^{33} + \alpha^{23}x^{34} + \alpha^{25}x^{40}$
11	$x^9 + \alpha^4(x^{10} + x^{18}) + \alpha^9(x^{12} + x^{20} + x^{40})$
12	$\alpha^{52}x^3 + \alpha^{47}x^5 + \alpha x^6 + \alpha^9 x^9 + \alpha^{44}x^{12} + \alpha^{47}x^{33} + \alpha^{10}x^{34} + \alpha^{33}x^{40}$
13	$\alpha(x^6 + x^{10} + x^{24} + x^{33}) + x^9 + \alpha^4 x^{17}$

Theorem 4. *Let n be even and $k = n/2$. Let P be a permutation polynomial of $\mathbb{F}_{2^n}$ with coefficients in $\mathbb{F}_{2^k}$, $P \in \mathbb{F}_{2^k}[x]$, then P cannot be APN.*

In particular, a polynomial with coefficients in a sub-field $\mathbb{F}_2$, is not APN. Until 2009, this result was supporting the idea that there are no APN permutations in $\mathbb{Z}_2^n$ for even n. Nevertheless, the challenge of finding APN functions or APN permutations have interested many researchers. In 2006, in a meeting held in the Banff International research station, John F. Dillon presented a list of APN quadratic functions in dimension 6 (see Table 3). A similar list for dimension 7 and 8 can be found in [48].

As the exhaustive search is not possible in dimension 6, most of the research has been dedicated to quadratic polynomials. Nevertheless, in [49], Edel and Pott proposed a new technique, the switching method, for deriving new APN functions. From Theorem 9 of [49], an instantiation of the switching method can be derived. It proceeds as follows:

- Select an APN function $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$, and $u \in \mathbb{Z}_2^n$.
- Create a Boolean function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ by solving a system of linear equations verifying that $f(x) + f(x + a) + f(y) + f(y + a) = 0$ if $F(x) + F(x + a) + F(y) + F(y + a) = 0$, for all $(x, y, a) \in (\mathbb{Z}_2^n)^3$.
- Then according to Theorem 9 of [49], the function $G(X) = F(x) + uf(x)$ is APN.

Application of this method in dimensions 6 and 8 has produced new non-quadratic APN functions. For instance, from the function number 6 of the Banff list (Table 3), the following function defined in $\mathbb{F}_{2^6}[x]$ has been derived

$$F(x) = x^3 + \alpha^{17}\left(x^{17} + x^{18} + x^{20} + x^{24}\right) + \alpha^{18}x^9 + \alpha^{36}x^{18} + u^9x^{36} + x^{21} + x^{42} + \mathbf{Tr}\left(\alpha^{27}x + \alpha^{52}x^3 + \alpha^6x^5 + \alpha^{19}x^7 + \alpha^{28}x^{11} + \alpha^2x^{13}\right),$$

where α is a root of the primitive polynomial $x \mapsto x^6 + x^4 + x^3 + x + 1$. This function is a non-quadratic APN function in dimension 6. Nevertheless, this method has not been successful at finding APN permutations in even dimension. Recently, Dillon et al. [4] provided an example of an APN permutation in $\mathbb{F}_{2^n}$ with even n. This function is defined over $\mathbb{F}_{2^6}$ as follows:

$$\begin{aligned} F(x) = {} & w^{45}x^{60} + w^{41}x^{58} + w^{43}x^{57} + w^{4}x^{56} + w^{50}x^{54} + w^{20}x^{53} + w^{45}x^{52} + w^{20}x^{51} \\ & + w^{23}x^{50} + w^{36}x^{49} + w^{56}x^{48} + w^{21}x^{46} + w^{5}x^{45} + w^{21}x^{44} + w^{28}x^{43} + w^{3}x^{42} \\ & + w^{59}x^{41} + w^{58}x^{40} + w^{57}x^{39} + w^{53}x^{38} + w^{37}x^{37} + w^{40}x^{36} + w^{18}x^{35} + w^{41}x^{34} \\ & + w^{54}x^{33} + w^{3}x^{32} + w^{49}x^{30} + w^{41}x^{29} + w^{42}x^{28} + w^{50}x^{27} + w^{53}x^{26} + w^{58}x^{25} \\ & + w^{9}x^{24} + x^{23} + w^{28}x^{22} + w^{3}x^{21} + w^{21}x^{20} + w^{52}x^{19} + w^{60}x^{17} + w^{59}x^{16} \\ & + w^{10}x^{15} + w^{42}x^{13} + w^{8}x^{12} + w^{35}x^{11} + w^{44}x^{10} + w^{45}x^{8} + w^{8}x^{7} + w^{61}x^{6} \\ & + w^{59}x^{5} + w^{20}x^{4} + w^{12}x^{3} + w^{37}x^{2} + w^{2}x, \end{aligned} \tag{3}$$

where $w = \alpha^{-2}$ and α is a root of the primitive polynomial $x \mapsto x^6+x^4+x^3+x+1$. This function has algebraic degree 4 and is CCZ-equivalent to a known non-bijective quadratic APN function $h(x) = x^3 + x^{10} + \alpha x^{24}$ which is the function number 5 in Table 3.

The idea of using CCZ-equivalence for finding APN permutations in $\mathbb{F}_{2^6}$ has so far been successful in producing, up to CCZ-equivalence, only one APN permutation. More generally, this function is the only known APN permutation of any $\mathbb{F}_{2^n}$, n even, and the question of the existence of APN permutations for even n, $n > 6$, is still pending [4].

Some of the other methods proposed for constructing APN functions use AB functions, others make uses of APN functions in smaller or larger fields as starting points. These methods do not always guarantee that the resulting function is strictly APN, but just has a small differential uniformity. In Section 6 we discuss some other methods used by cryptographers to produce functions with small differential uniformity.

There has also been some works towards enumerating the number of CCZ-inequivalent APN functions over $\mathbb{Z}_2^n$ when $n \leq 8$. For $n = 7$, the latest known result [50] shows that there is at least 470, CCZ-inequivalent APN functions over $\mathbb{Z}_2^7$. These results illustrate the difficulty of finding such functions for large n in characteristic 2. Similar works for PN functions, can be found, e.g. [51], in characteristic p, when the dimension of the vectorial space is small.

5. The exponential and logarithm functions

In the context of the real functions, the exponential function is distinguished by the property of being its own derivative. A similar property holds for the discrete exponential function in finite fields.

Let α be a primitive element in a finite field $\mathbb{F}_{p^n}$. Then the discrete exponential function G in $\mathbb{F}_{p^n}$ is defined as

$$G : \mathbb{Z}_{p^n-1} \to \mathbb{F}_{p^n}, \quad x \mapsto \alpha^x.$$

Then, for any difference $a \neq 0$ in $\mathbb{Z}_{p^n-1}$, we have

$$G(x+a) - G(x) = \alpha^{x+a} - \alpha^x = \alpha^x\big(\alpha^a - 1\big) = \big(\alpha^a - 1\big)G(x), \quad \text{for all } x \in \mathbb{Z}_{p^n-1}.$$

As G is one-to-one, also all its derivatives are one-to-one. Thus, the discrete exponential function G is a PN function from the Abelian group $(\mathbb{Z}_{p^n-1}, +)$ to the Abelian group $(\mathbb{F}_{p^n}, +)$. But it is not bijective. It can be extended to a bijective function in different ways. The first approach to be discussed next is to add one element to the domain of definition. The second approach consists of removing one element from the image space of the function and is presented at the end of this section.

5.1. The bijective discrete exponential and logarithm functions

A bijective extension of G is obtained by adding one element to the domain of G and defining it as $\mathbb{Z}_{p^n}$, and then setting $F(x) = G(x)$, for $x \neq p^n - 1$ and $F(p^n - 1) = 0$. We call F the bijective discrete exponential function in $\mathbb{F}_{p^n}$. Its inverse is called the bijective discrete logarithm function in $\mathbb{F}_{p^n}$. These functions are promising candidates for S-boxes in cipher constructions and allow analysis of some other nonlinear functions. In particular, it was shown in [52] that the nonlinearity properties of the key multiplication operation in the block cipher IDEA [53] are related to nonlinearity properties of the bijective discrete exponential and logarithm functions with $p = 2^{16} + 1$ and $n = 1$.

The resulting function F is at most differentially 4-uniform, which we can verify as follows. Given $a \in \mathbb{Z}_{p^n}$, $a \neq 0$, and $b \in \mathbb{F}_{p^n}$ we split the domain into four parts as follows to be able to use the PN property of the function G.

1. $0 \leq x < p^n - 1 - a$: solve $G(x+a) - G(x) = b$.
2. $x = p^n - 1 - a$: gives one solution if $b = -G(p^n - 1 - a)$.
3. $p^n - 1 - a < x < p^n - 1$: solve $G(x+a-1) - G(x) = b$.
4. $x = p^n - 1$: one solution if $b = G(a-1)$.

Each of the four cases gives at most one solution. Note that if $a = p^n - 1$ the first case is empty.

For cryptographic purposes the size of the domain is often selected as being a power of 2. Since $p = 2^{2^4} + 1$ is the largest known Fermat prime, there are no larger primes of this form of practical size. Another approach is to consider $\mathbb{F}_{2^n}$ and extend the integer domain from $\mathbb{Z}_{2^n-1}$ to $\mathbb{Z}_{2^n}$. By identifying the set $\mathbb{Z}_{2^n}$ with the linear space $\mathbb{Z}_2^n$ we can investigate the nonlinearity properties of the discrete exponential and logarithm

functions considered as bijective functions in $\mathbb{Z}_2^n$. Nina Brandstätter, Tanja Lange and Arne Winterhof [54] were the first to investigate the nonlinearity of the least significant bit of the bijective discrete logarithm function. Let us denote by f the least significant bit of a bijective discrete logarithm function F in $\mathbb{Z}_2^n$. Then it was proved in [54] that

$$\mathcal{L}(f) \leq 2n2^{\frac{n}{2}}.$$

Later also the nonlinearity of other components of the bijective discrete logarithm has been investigated [55,56], and the following bound

$$\mathcal{L}(f_\lambda) \leq \frac{8}{\pi^2}\left(2^{k+1}+1\right)\left(\ln\left(2^n-1\right)+2\right)2^{\frac{n}{2}}$$

was obtained in [56]. Here k is the position of the most significant non-zero bit of the linear mask λ that defines the component $f_\lambda(x) = \langle \lambda, F(x)\rangle$ of F. It was also shown that the linearity $\mathcal{L}(f_\lambda)$ of a component f_λ is invariant under rotation of the bit vector λ. Hence the least value k over all rotations of λ can be used. Recently, this bound was slightly improved in [57,58].

For very sparse masks, the upperbound can be improved as it was shown in [52] to give the following bound

$$\mathcal{L}(f_\lambda) \leq 1 + n^{w_H(\lambda)}2^{\frac{n}{2}},$$

where the Hamming weight of λ is denoted by $w_H(\lambda)$. This bound holds for $n \geq 5$. The proof makes use of the fact that, by the rotation property, all one-bit masks have the same linearity upperbound as the least significant bit.

Some experiments have been conducted to validate these bounds [56]. By these experiments the nonlinearity of F grows with n at about the same rate as the nonlinearity of a single coordinate function. This indicates that the upperbound of linearity that increases exponentially with the length or with the Hamming weight of the masks is far too large. It remains an open problem how to make these bounds tighter to obtain better estimates of the apparently high nonlinearity of the bijective discrete logarithm function.

5.2. The Welch–Costas functions

We start again from the PN discrete exponential function G defined in the beginning of this section and fix $n = 1$. This function is PN from the Abelian group $(\mathbb{Z}_{p-1}, +)$ to the Abelian group $(\mathbb{F}_p, +)$. To make it bijective, we subtract 1 from all values of $G(x)$ and then remove the element $p-1$ from the original image set $\mathbb{F}_p = \mathbb{Z}_p$ to replace it by $\mathbb{Z}_{p-1}$. The resulting function is called the exponential Welch–Costas (EWC) function and we denote it by F in this subsection. Then F is a bijection in $\mathbb{Z}_{p-1}$ and

$$F(x) = \left(\alpha^x \bmod p\right) - 1, \quad \text{for all } x \in \mathbb{Z}_{p-1}.$$

Its inverse function is called the logarithmic Welch–Costas (LWC) function. These functions are APN [59].

These functions have found applications in cryptography. For example, the SAFER family of block ciphers [60,61] exploits EWC and LWC functions with parameters $p = 2^8 + 1$, $\alpha = 45$ and $n = 1$.

The linearities of the EWC function $F : \mathbb{Z}_{p-1} \to \mathbb{Z}_{p-1}$ and the corresponding discrete exponential function $G : \mathbb{Z}_{p-1} \to \mathbb{Z}_p$ have been studied in the context of the generalized definition of linearity in Abelian groups. By definition,

$$\mathcal{L}(F) = \max_{0 \le a < p-1, 0 < b < p-1} \left| \sum_{x \in \mathbb{Z}_{p-1}} w_{p-1}^{bF(x)} \overline{w}_{p-1}^{ax} \right|, \quad \text{and}$$

$$\mathcal{L}(G) = \max_{0 \le a < p-1, 0 < b < p} \left| \sum_{x \in \mathbb{Z}_{p-1}} w_{p}^{bG(x)} \overline{w}_{p-1}^{ax} \right|.$$

Then $\sqrt{p-1} \le |\mathcal{L}(F)| \le p - 1$ and the same holds true for G. Moreover, it is quite straightforward to show using the PN property G and some ordinary properties of Fourier transforms that $|\mathcal{L}(G)| \le \sqrt{p}$ [62]. For the EWC function F the situation is not equally straightforward. Based on extensive experiments it was conjectured that the linearity of F is bounded from above by a constant times $p^{\frac{1}{2}+\epsilon}$, where ϵ is some small constant. This conjecture was proved to hold by Risto Hakala in [63]. Moreover, Hakala proved with the help of a method from Louis J. Mordell's last paper (1972) that the following inequality holds

$$\left|\mathcal{L}(F)\right| \le \left(\frac{2}{\pi} \ln p + 4\right) \sqrt{p}$$

and hence gave an upperbound that is asymptotically strictly less than the bound originally conjectured by Konstantinos Drakakis, Verónica Requena and Gary McGuire [62].

To summarize, let us note that the square root of the size of the domain is the dominating factor of all linearity bounds of the different variants of the discrete exponential and logarithmic functions studied in this section. This factor is multiplied by a factor that is only of logarithmic of the size of the domain of the function. It is an interesting question, how small the logarithmic factor can be made in particular in the case of the vectorial Boolean function based on the bijective discrete logarithm function.

6. Applications in cryptography

In this section, we discuss the use of PN and APN functions in the symmetric key cryptographic algorithms such as block ciphers and hash functions. Such constructions are usually defined iteratively, by repeating interleaving nonlinear and linear layers. The nonlinear layers ensure the confusion through a parameterized permutation. The linear layers ensure the diffusion through the cipher.

A nonlinear layer is typically defined as a parallel application of S-boxes. We start by discussing S-boxes defined on extension of $\mathbb{Z}_2$. Two main types of iterative structures can be distinguished, the Substitution Permutation Network (SPN) and the Feistel Network with all its generalizations. The Feistel structure guarantees bijectivity independently of the bijectivity of the round functions that are used to modify the state. Nevertheless, as summarized in [64], these round functions should be chosen with care since some attacks can take advantage of non-injective or non-surjective S-boxes. For SPN structures the S-boxes must be bijective.

6.1. Computation of the difference table

One of the obstacles encountered in the search of APN functions is the verification of this property which algorithmically requires the computation of the values of a half of the derivatives. More precisely [36], given a hyperplane $\mathcal{H}$, if for all $a \in \mathcal{H}$ and for all $b \in \mathbb{Z}_p^n$ we have $\delta(a,b) \leq 2$, then the function is APN. The time for such computation is 2^{2n-1}. For particular polynomials, such as monomials over $\mathbb{F}_{p^n}$ this computation can be reduced to the study of only one derivative as explained in Section 4.1. But also other special classes of functions have been identified where fast verification of the APN property is possible. One such class is composed of polynomials with coefficients in a sub-field [65]. In particular, for a polynomial $P : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ which has coefficients in $\mathbb{F}_2$ we have $\delta(a,b) = \delta(a^{2^i}, b^{2^i})$, for all $0 < i < n$. In particular, for such polynomials we have that $\delta(1,b) = \delta(1,b^{2^i})$, for all $0 < i < n$. Using this property one can prove that, if n is odd and for all $b \in \mathbb{F}_{2^n}$ $\delta(1,b) = 0$ or 2, then $\delta(1,1) = 2$ if $P(0) = 0$. Notice that similar observations can be used to prove the non-existence of APN permutation polynomial P of $\mathbb{F}_{2^n}$ when $n = 2^k$ with $P \in \mathbb{F}_2[x]$, $P(0) = 0$. Indeed from the cyclotomic class we deduce that for such polynomials $P(1) = 1$, and $\sum_x P(x) = \sum_{b,\delta(1,b)=2} b = 0$. From the number of elements in the cyclotomic class of order greater than 2 we deduce a contradiction.

For usage in cryptography, the differential properties of a function $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$, are often presented in a two dimensional array called the difference distribution table (DDT) denoted by T_F. The entries of T_F are the values of $\delta(a,b)$ for all $a \in \mathbb{Z}_p^n$ and $b \in \mathbb{Z}_p^m$. For an APN function, all entries in this table for $a \neq 0$ are equal to either 0 or 2. An example of a DDT of a permutation over $\mathbb{Z}_2^3$ is given in Table 4. If F is a permutation, $\delta(a,0) = 0$ for all $a \neq 0$.

Knowing the representation, the properties summarized in the previous paragraph can be visually observed on this table. For instance, for a monomial function, all lines of the DDT are permutations of each other.

More generally, the complete knowledge of the DDT of an APN function allows us to recover the values of the function. Indeed, from the fact that $\delta(a,b) = 2$ it follows that there exists a unique pair $(x, x+a)$ such that $f(x+a) + f(x) = b$. Writing these partial equations allows us to determine one of the corresponding functions in the class of the equivalence. For instance from Table 4, we have that $f(0) + f(1) = 1, 3, 5$ or 7, assuming

Table 4
A difference distribution table of an APN permutation defined over $\mathbb{Z}_2^3$.

a/b	0	1	2	3	4	5	6	7
0	8	0	0	0	0	0	0	0
1	0	2	0	2	0	2	0	2
2	0	0	0	0	2	2	2	2
3	0	2	0	2	2	0	2	0
4	0	0	2	2	0	0	2	2
5	0	2	2	0	0	2	2	0
6	0	0	2	2	2	2	0	0
7	0	2	2	0	2	0	0	2

that $f(0) = 0$ gives us only 4 values for $f(1)$. The number of derived partial equations are enough to determine one of the APN functions in the class of equivalence.

The difference tables T_F and T_G of two affine equivalent functions F and G are permutation of each others. In particular, if $G = F(L_1(x))$ with L_1 an affine permutation, we have $\delta_G(a, b) = \delta_F(L_1(a), b)$, for all a and b, meaning that the rows of T_G are the same as the rows of T_F, but they appear in a permuted order. Similarly, if $G = L_2(F(x))$ with L_2 an affine permutation, we have $\delta_G(a, b) = \delta_F(a, L_2^{-1}(b))$ and the columns of T_G are the same as the ones of T_F but they appear in a permuted order. The extended affine equivalence is less visual, since if $G(x) = F(x) + L(x)$ then $\delta_G(a, b) = \delta_F(a, b + L(a))$.

6.2. Provable security

Differential cryptanalysis is a statistical method of cryptanalysis which was developed for attacking the NSA block cipher DES [3]. After its introduction, Kaisa Nyberg and Lars R. Knudsen studied the possibility of designing a DES-type block cipher with provable resistance against differential attacks. In [66,67], they provide bounds on the probability of a 4-round DES-like cipher. The number of rounds can be reduced to three if the DES-like cipher has a bijective round function. The bounds were later improved by Kazumaro Aoki [68]. This result shows a close relationship between the differential property of the round function and the resistance of the cipher to differential cryptanalysis. Based of this idea Nyberg and Knudsen proposed a 64-bit DES-like cipher which makes use of the quadratic monomial $x \mapsto x^3$ over the field $\mathbb{Z}_2^{33}$. The designers called the cipher by the name CRADIC standing for "Cipher Resistant Against Differential Cryptanalysis", while in the literature it is better known as the $\mathcal{KN}$ cipher. It was soon broken by higher order differential attacks [69] that make use of the low algebraic degree of the x^3 function. A similar attack can be conducted if instead of the x^3 function the inverse function x^{-1} was used. Recently, Christina Boura and Anne Canteaut [70] followed the suggestion given in [67] and investigated the inverses of the quadratic functions x^{2^k+1} as the round function for a CRADIC-like cipher and show positive results about its potential resistance against algebraic attacks.

In [71], Mitsuru Matsui generalized the notion of provable security for unbalanced Feistel structures and other modified structures that were later to be known as MISTY structures. Since in 1996, APN permutations over $\mathbb{Z}_2^n$ with n even were believed not to

exist, Matsui split the state into two branches with odd number of bits thus allowing to use APN permutations. He showed that in this way the security against differential cryptanalysis is better than what one would obtain by [66] using branches of equal size and an even number of bits. In 1997, based on this idea, Matsui proposed the MISTY ciphers [72]. These ciphers are imbrications of Feistel networks defined over 64 bits, 32 bits and 16 bits. The innermost structure of these ciphers is an unbalanced modified Feistel network which uses as round functions two nonlinear permutations of 7 and 9 bits. To limit the hardware print, the 7-bit S-box was chosen among linear transforms of monomials with exponents of weight three. The resulting exponent is a Welsh exponent. It is interesting to note that while the infinite family of monomials with Welsh exponent $d = 2^m + 3$, defined over $\mathbb{F}_{2^n}$, for $n = 2m+1$ odd, was not proved APN until in 1999 [34, 35] (see Section 4.1), such an APN was relatively easy to find in this dimension by computer search. The algebraic degree of this function is 3. A quadratic function was chosen for the 9-bit S-box. Using similar design principles, new different S-boxes have been selected for the KASUMI variant of MISTY [73].

Later similar works have been conducted to prove the security of some generalized Feistel or SPN constructions against linear and differential cryptanalysis. These results show that S-boxes with good differential properties do not alone guarantee the resistance against differential cryptanalysis. Indeed, one should also take into consideration the number of S-boxes that are activated when injecting a difference between two plaintexts. To guarantee that this number is sufficiently large, the notion of branch number [74] has been introduced to capture this behavior of the diffusion layer and used in providing security claims for block-oriented primitives, such as the AES.

In the AES, certain linear and non-linear functional components can be composed together to a larger functional entity called a Super S-box to produced improved cryptanalysis on the AES [75–77], AES-like block ciphers and hash functions. In [78], Anne Canteaut and Joëlle Roué studied differential properties of combinations of a linear diffusion layer and an S-box layer over two rounds and derived criteria for choosing functions for these layers in an optimal way.

6.3. S-boxes in practice

As recalled in Section 3.3, linear and differential properties of a function are preserved by the affine and EA-equivalence as well as by the CCZ-equivalence. While the first two equivalences preserve also algebraic degree of the function, the implementation cost of the nonlinear layer is different depending of the chosen function in the equivalence set. To select the best S-box in regard to particular constraints related to the implementation environment such as speed or physical attacks, it is common to check the cost of a set of functions in the affine equivalence set.

Usage of the affine equivalence to limit the implementation cost can be found already in [71] for the design of the MISTY block cipher. Later Leander and Poschmann [79] derived the list of all the differentially 4-uniform permutations of $\mathbb{Z}_2^4$. Explanation on

how to reduce the exhaustive search using the EA-equivalence is then provided. This approach is not effective for larger dimension and in [80], another method to find all APN permutations up to dimension 5 is described.

In more recent designs, affine equivalence has been used to find S-boxes with low latency [81] and S-boxes easier to protect against side channel attacks [64,82,83]. In these recent works, no concrete methods are proposed to find functions which meet the requirements and the functions are found by randomly searching for the best ones among the equivalent functions. This approach allows search while preserving good differential and nonlinearity properties. Another approach is to focus at limiting the number of instructions needed for the implementation of S-boxes. For instance in [84], the authors concentrate on 4-bit S-boxes and show how to find a permutation which is differentially 4-uniform, using only 9 instructions. This S-box is good also with respect to other nonlinearity properties.

In general, while APN functions are mathematically interesting, their usefulness for cryptography is limited. Next we will discuss some well-known designs, the choices made when selecting the S-boxes, and the resulting security of these constructions.

AES and locally-APN functions. When in 1998, the NIST has launched a competition with the goal of establishing a new block cipher standard, the list of known APN permutations was still relatively small and no APN permutation in even dimensions was known. Due to this reason and some designs restrictions, differentially 4-uniform S-boxes were considered in the design of some ciphers such as E2, the ancestor of Camellia [85], AES [29] and also in more recent designs (e.g. [86,81,87]).

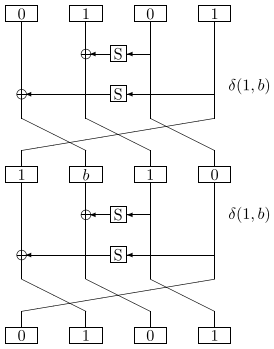
In the design of the AES block cipher, Joan Daemen and Vincent Rijmen defined the S-box as an affine transformation of the inverse function in the finite field $\mathbb{F}_{2^8}$. This permutation, which is differentially 4-uniform [8], achieves the best known differential uniformity in this space of dimension 8. In particular, the derivatives of this function are such that $\delta(1,1) = 4$ and for all $b \neq 1$, $\delta(1,b) \leq 2$. In [88], this function gave rise to the notion of locally-APN monomials.

Definition 6. (See [88].) Let $F_d : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$, $x \mapsto x^d$ be a power function. The function F is called locally-APN if for all $b \neq 0, 1$ we have $\delta(1,b) \leq 2$.

There is some reason to believe that the locally-APN S-boxes might provide better resistance to differential cryptanalysis than other functions with the same differential uniformity. This will be demonstrated next using a cryptographic toy example. Let us first recall some notation related to the differential spectrum of a function. Generalizing the notation of [89], for a fixed $a \in \mathbb{Z}_q^n$, let us denote by ω_i^a the quantity

$$\omega_i^a = \left|\left\{b \mid \delta(a,b) = i\right\}\right|.$$

The differential spectrum of F is then defined as the sequence of all ω_i^a, $i \leq \delta(F)$. In characteristic 2, we have $\omega_i^a \neq 0$ only for even values of i. If F is a monomial function, the



S-box	Differential spectrum			
S_i	ω_4^1	ω_2^1	ω_0^1	p_i
S_0	0	2^{n-1}	2^{n-1}	$\frac{2^{n+1}}{2^{2n}}$
S_1	1	$2^{n-1}-2$	$2^{n-1}+1$	$\frac{2^{n+1}+8}{2^{2n}}$
S_2	2^{n-2}	0	2^n-2^{n-2}	$\frac{2^{n+2}}{2^{2n}}$

Fig. 1. Example of differential computation for a generalized Feistel of [91].

knowledge of ω_i^1 is enough to determine the whole spectrum. The notion of locally-APN monomial has later been generalized to the notion of locally differentially λ-uniform in [90].

Given an S-box $S = S_i$, the probability $p = p_i$ of the differential defined in Fig. 1 can be computed as the sum of the probabilities of all differential trails as

$$p = \sum_b \left(2^{-n}\delta(1,b)\right)^2.$$

For the APN S-box S_0 we have that $p_0 = 2^{-n+1}$. For the differentially 4-uniform locally APN S-box S_1 we have $p_1 = 2^{-n+1} - 2^{3-2n}$. The differentially 4-uniform S-box S_2 has a 2-valued differential spectrum and $p_2 = 2^{2-n}$. This example demonstrates that a locally APN S-box can give smaller differential probabilities than other differentially 4-uniform S-boxes.

It was shown in [79] that among the permutations of $\mathbb{Z}_2^4$ the only equivalence class comprising of differentially 4-uniform and locally-APN functions corresponds to the one of the inverse function. Up to equivalence, the inverse function is the only known differentially 4-uniform locally-APN function in $\mathbb{Z}_2^n$, for even n.

Because of its very good differential and linear properties, many other cryptographic primitives are using functions affine equivalent to the inverse function for confusion. In the beginning of this century [92] it has been thought from the point of view of algebraic cryptanalysis that the property $x \times x^{-1} = 1$ is suspect. Nevertheless, as shown in [93], the XSL algorithm exploiting this property did not work as expected. Up to now, no attack on the AES exploiting this property has been found.

Keccak and quadratic functions. Among other cryptographic standards, we describe briefly some properties of the winner Keccak [94] of the SHA3 competition organized by the NIST. The design of this hash function is based on a sponge construction using an internal permutation of 800 or 1600 bits. Among the different operations which compose this permutation, the nonlinear layer is defined as a parallel application of a 5-bit permutation S-box $\chi : \mathbb{Z}_2^5 \mapsto \mathbb{Z}_2^5$ defined coordinate by coordinate as follows:

$$\chi(x_0, x_1, x_2, x_3, x_4) = \begin{pmatrix} x_0 + x_2 + x_1x_2 \\ x_1 + x_3 + x_2x_3 \\ x_2 + x_4 + x_3x_4 \\ x_3 + x_0 + x_4x_0 \\ x_4 + x_1 + x_0x_1 \end{pmatrix}. \tag{4}$$

All components of this S-box are quadratic. By analyzing the Keccak function, one can see that the values of $\delta(a, b)$ are related to the Hamming weight of a. In particular, $\delta(a, b) \in \{0, 8\}$ for a of weight 1, $\delta(a, b) \in \{0, 4\}$ for a of weight 2, and finally, $\delta(a, b) \in \{0, 2\}$ for a of weight 4 or 5. If a is of weight 3, we have $\delta(a, b) \in \{0, 4\}$ or $\delta(a, b) \in \{0, 2\}$ depending on the distribution of the bits 1 in the binary representation of a.

In general, the derivatives of quadratic functions are 2^t-to-1. For the components of the Keccak function, t varies from 1 to 3. The Keccak function (4) serves as a demonstration of the fact that even an S-box with non-APN quadratic components can be used to obtain a solid cryptographic primitive. Thanks to the strong diffusion, only few rounds of the Keccak permutation can be distinguished from random using a differential distinguisher.

FIDES and the APN permutation in dimension 6. It was not until in 2013 when the remarkable APN permutation over $\mathbb{F}_{2^6}$ found its use in a cryptographic design. The permutation of the FIDES [83] authenticated encryption primitive is composed of 32 parallel applications of this APN permutation. The polynomial representation of it being rather complex to evaluate, implementation as a truth table is recommended.

To evaluate the resistance against other cryptanalytic attacks, also other cryptographic properties of this APN permutation are relevant. From the polynomial provided in (3), one can check that the algebraic degree of this function is 4. As recalled in Section 3.1, the minimal known linearity of a permutation defined over $\mathbb{F}_{2^n}$ for n even is $2^{n/2+1}$ and is equal to 16 in dimension 6. This value which is reached by the inverse function is also reached by this APN permutation. As the inverse function is locally-APN and has algebraic degree 5, differential and linear attacks on this cipher might be similar independently of the choice of S-box EA-equivalent to the inverse function or of (3). Notice that among others, FIDES is an example of construction where using APN S-boxes does not guarantee its security. For instance in [95], using a guess-and-determine method, an attack on the full primitive is proposed. This attack relies on the weakness of the diffusion provided by the linear layer. The differential properties of the S-box do not play any role in this attack.

6.4. Large S-boxes

For implementation reasons, S-boxes are often defined as permutation of 4 or 8-bits. Since the probability of a differential is related to the size of the S-box, and is defined by $P[a \to b] = \frac{\delta(a,b)}{2^n}$, it would be more beneficial to use large S-boxes. After the design of the $\mathcal{KN}$ cipher, where the idea of a large S-box appears for the first time, also other works towards this direction have appeared. For instance, as 8-bit S-boxes are usually

better in term of differential, linear and algebraic properties, than 4-bit S-boxes, but are more costly to implement, cryptographers have been investigating methods to construct large S-boxes from smaller ones. Such an idea was taken into consideration for instance in the design of ICEBERG [96] and ZORRO [82]. The resistance to differential and linear cryptanalysis is then measured in regard to the 8-bit S-box. Inspired by [97], the authors of PICARO explain in [64] why the selected S-box, defined over $\mathbb{F}_{2^8}$ and admitting presentation as a function of two bivariate polynomials in $\mathbb{F}_{2^4}$, is a suitable candidate for an efficient side channel resistant cipher. More discussion on such approaches can be found in [82]. While the main objective has been to limit the number of operations necessary for the implementation and to ensure a certain protection against side-channel attacks, the design of a function with small differential uniformity and linearity remains as another important goal. Therefore the direction of study proposed in [97] seems interesting for cryptography.

6.5. Exploiting the APN property of the S-boxes in cryptanalysis

Since the APN property implies exceptional and extreme behavior, it can in some situations lead to vulnerabilities of the cipher where APN components are being used. In [98], Anne Canteaut and Marìa Naya-Plasencia present an attack on the hash function MARACA proposed to the SHA3 competition and show that the higher the number of input differences which can lead to the same output difference is, the better the attack works. The quantity describing the power of the attack is denoted by ∇_F and is defined as

$$\nabla_F = \max_{b \in \mathbb{Z}_2^n} \#\mathcal{D}_F(b) \quad \text{where } \mathcal{D}_F(b) = \{a \mid x,\ F(a) + F(x+a) = b \text{ for some } x\}.$$

Given a permutation function F over $\mathbb{Z}_2^n$, a lower bound on ∇_F is $2^n/\Delta_F$. The functions with small differential uniformity are the ones with the largest ∇_F value.

Among the functions for which the value of ∇_F is fully determined we give the example of the quadratic monomials. Let $Q_t : x \mapsto x^{2^t+1}$ over $\mathbb{F}_{2^n}$ and let $s = \gcd(t, n) \geq 1$, one can show that for any a, b in $\mathbb{F}_{2^n}$, if x is a solution of $Q_t(x) + Q_t(x+a) = b$, the set of solutions of this equation is $x + a\mathbb{F}_{2^s}$. Meaning that for all a, b in $\mathbb{F}_{2^n}$, $\delta(a, b) \in \{0, 2^s\}$ and $\nabla_F = 2^{n-s}$.

The concept of two-valued differential spectrum of a vectorial Boolean function is studied in [89]. In particular, let Δ_F be the differential uniformity of a monomial function $F(x) = x^d$. This function is differentially two-valued if for all $b \in \mathbb{F}_{2^n}$, $\delta(1, b)$ is equal to 0 or Δ_F. In this case it can be shown [89] that Δ_F is a power of two. Some monomials with Kasami exponent have a two-valued differential spectrum.

Theorem 5. *(See [89].) Let $K_t : x \mapsto x^{2^{2t}-2^t+1}$ over $\mathbb{F}_{2^n}$. We assume that $n \neq 3t$ and $s = \gcd(n, t)$ with n/s odd. Then, for any $b \in \mathbb{F}_{2^n}$, $\delta(b) \in \{0, 2^s\}$.*

As noticed in [98], F is differentially Δ_F-uniform with two-valued differential spectrum if and only if $\nabla_F \cdot \Delta_F = 2^n$ holds.

The attack described in [98] does not take advantage of only the size of $\mathcal{D}_F(b)$, but also of the structure of these sets, and gives rise to the following generalization of crooked functions originally defined in [99].

Definition 7. (See [99,98].) A function $F : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is said to be crooked (resp. crooked of co-dimension $d \geq 1$) if for any non-zero $b \in \mathbb{Z}_2^n$, $\mathrm{Im}(D_a F)$ is an affine subspace of co-dimension 1 (resp. co-dimension d).

The function F is weakly crooked (resp. weakly crooked of co-dimension d) if $\mathrm{Im}(D_a F)$ is contained in an affine subspace of co-dimension 1 (resp. co-dimension d).

Crooked permutations are AB and therefore APN Boolean functions on an odd number of variables. In [67], all quadratic permutation monomials were proved crooked. In [98] it is conjectured that crooked permutations of any co-dimension are quadratic.

While AB and APN functions are related to the well known cryptanalytic methods, the work of Canteaut and Naya-Plasencia shows nevertheless that crooked functions might be avoided in some cryptographic applications.

6.6. Other types of ciphers

Not all symmetric cryptographic primitives use S-boxes. Instead, nonlinearity can be ensured by the use of nonlinear operations such as modular addition. Examples of such primitives are the block ciphers TEA, XTEA, as well as the hash functions MD5, SHA-1, SHA-2, and the recent 3rd round SHA-3 candidates: SKEIN and BLAKE. For constructions like the ARX one, the analysis presented in this paper may seem useless since the confusion is ensured by simple nonlinear operations such as AND and modular additions.

More recently, the NSA has published two new block ciphers [100], SIMON and SPECK which are defined as generalized Feistel ciphers. The nonlinearity of SIMON is ensured by a simple AND. For a cipher of $2b$ bits, one can consider it as a Super-Sbox (see Section 6.4) of b bits which is far from being APN and can be proved differentially 2^{b-2}-uniform. For this cipher, the security against differential cryptanalysis is then ensured primarily by repeating this simple round function a large number of times. The nonlinearity of the $2b$-bit block cipher SPECK is ensured by a modular addition of b state-bits with the other b state bits. The diffusion being faster than the one of the SIMON cipher, less rounds are required.

7. Conclusion

The planar, bent and generalized bent functions have been known in the mathematical literature since the 1980s. In the late 1980s the importance of highly nonlinear functions

in cryptography was first discovered by Meier and Staffelbach from the point of view of correlation attacks on stream ciphers, and later by Nyberg in the early 1990s after the introduction of the differential cryptanalysis method. The APN functions, which have the optimal properties for offering resistance against differential cryptanalysis, have since then been an object of intensive study by many mathematicians, and achieved a lot of interest for its own sake.

The goal of this brief survey has been to describe the state of the art of the APN and other closely related highly nonlinear functions from the point of view of a cryptographer. Still in late 1990s the designers of symmetric key cryptographic algorithms considered it of paramount importance to choose the nonlinear building block of their ciphers to have the best possible nonlinearity properties. In his hardware oriented MISTY designs, Matsui split the 16-bit state into two odd parts of different length to be able to use APN permutations. The designers Daemen and Rijmen of the AES algorithm, which was targeted for implementation also in software, could not go that far, and had to make a suboptimal choice which was the differentially 4-uniform inverse function.

Since then the cryptographic research community has gained a lot of new experience and understanding about the roles the different components have in providing cryptographic security. In this survey, we discussed FIDES, an example of recent designs with optimally nonlinear components, which was immediately broken due to its weaknesses in the linear component. The attack on MARACA exploits the observation that the evenly distributed differences, which make the differential probabilities small, make it possible to get a certain output difference from a maximum number of different input differences. On the other hand, several examples of very strong cryptographic designs are known that do not have optimal or even close to optimal nonlinearity in their components but achieve their strength by a large number of iterations.

The optimal or close to optimal nonlinearity and differential uniformity can offer one foundation for achieving solid arguments in favor of cryptographic strength. But it cannot be claimed sufficient and neither does it seem necessary. Still, the mathematical analysis of cryptographic building blocks remains valuable. In particular, new results on mathematical constructions of functions that can be implemented efficiently and offer protection to some different attacks simultaneously would be appreciated by cryptographers.

Acknowledgments

The authors wish to thank the anonymous reviewers for correcting many errors in the manuscript and giving insightful comments for improving the presentation of the paper.

References

[1] J.F. Dillon, Elementary Hadamard difference sets, in: Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory and Computing, Boca Raton, Florida, 1975, pp. 237–249.
[2] S. Murphy, The cryptanalysis of FEAL-4 with 20 chosen plaintexts, J. Cryptol. 2 (3) (1990) 145–154.

[3] E. Biham, A. Shamir, Differential cryptanalysis of DES-like cryptosystems, in: A. Menezes, S.A. Vanstone (Eds.), CRYPTO, in: Lect. Notes Comput. Sci., vol. 537, Springer, 1990, pp. 2–21.
[4] K.A. Browning, J.F. Dillon, M.T. McQuistan, A.J. Wolfe, An APN permutation in dimension six, in: Finite Fields: Theory and Applications, in: Contemp. Math., vol. 518, Amer. Math. Soc., 2010, pp. 33–42.
[5] P. Charpin, PN and APN functions (special functions over finite fields), in: G.L. Mullen, D. Panario (Eds.), Handbook of Finite Fields, in: Discrete Math. Ser., Chapman and Hall/CRC, 2013, Ch. 9.2.
[6] C. Carlet, Vectorial Boolean functions for cryptography, in: Y. Crama, P.L. Hammer (Eds.), Boolean Models and Methods in Mathematics, Computer Science, and Engineering, Cambridge University Press, 2010, pp. 398–469.
[7] P. Dembowski, T.G. Ostrom, Planes of order n with collineation groups of order n^2, Math. Z. 103 (1968) 239–258.
[8] K. Nyberg, Differentially uniform mappings for cryptography, in: T. Helleseth (Ed.), EUROCRYPT'93, in: Lect. Notes Comput. Sci., vol. 765, Springer, 1993, pp. 55–64.
[9] W. Meier, O. Staffelbach, Nonlinearity criteria for cryptographic functions, in: J.-J. Quisquater, J. Vandewalle (Eds.), EUROCRYPT'89, in: Lect. Notes Comput. Sci., vol. 434, Springer, 1989, pp. 549–562.
[10] O.S. Rothaus, On "bent" functions, J. Comb. Theory, Ser. A 20 (1976) 300–305.
[11] K. Nyberg, Perfect nonlinear S-boxes, in: D.W. Davies (Ed.), EUROCRYPT'91, in: Lect. Notes Comput. Sci., vol. 547, Springer, 1991, pp. 378–386.
[12] P. Kumar, R. Scholtz, L. Welch, Generalized bent functions and their properties, J. Comb. Theory, Ser. A 40 (1) (1985) 90–107.
[13] Y. Zhou, $(2^n, 2^n, 2^n, 1)$-Relative difference sets and their representations, J. Comb. Des. 21 (12) (2013) 563–584, http://dx.doi.org/10.1002/jcd.21349.
[14] F. Chabaud, S. Vaudenay, Links between differential and linear cryptanalysis, in: A.D. Santis (Ed.), EUROCRYPT'94, in: Lect. Notes Comput. Sci., vol. 950, Springer, 1994, pp. 356–365.
[15] A. Canteaut, P. Charpin, H. Dobbertin, Weight divisibility of cyclic codes, highly nonlinear functions on $\mathbb{F}_{2^m}$, and crosscorrelation of maximum-length sequences, SIAM J. Discrete Math. 13 (1) (2000) 105–138.
[16] D. Sarwate, M. Pursley, Crosscorrelation properties of pseudorandom and related sequences, Proc. IEEE 68 (5) (2005) 593–619.
[17] C. Blondeau, K. Nyberg, New links between differential and linear cryptanalysis, in: T. Johansson, P.Q. Nguyen (Eds.), EUROCRYPT'13, in: Lect. Notes Comput. Sci., vol. 7881, Springer, 2013, pp. 388–404.
[18] C. Blondeau, K. Nyberg, Links between truncated differential and multidimensional linear properties of block ciphers and underlying attack complexities, in: E. Oswald, P.Q. Nguyen (Eds.), EUROCRYPT'14, vol. 8441, Springer-Verlag, 2014.
[19] C. Carlet, C. Ding, Highly nonlinear mappings, J. Complex. 20 (2–3) (2004) 205–244.
[20] O.A. Logachev, A.A. Sal'nikov, V.V. Yashchenko, Bent functions on a finite abelian group, Diskretn. Mat. 9 (4) (1997) 3–20.
[21] C. Carlet, P. Charpin, V. Zinoviev, Codes, bent functions and permutations suitable for DES-like cryptosystems, Des. Codes Cryptogr. 15 (2) (1998) 125–156.
[22] C. Li, L. Qu, S. Ling, On the covering structures of two classes of linear codes from perfect nonlinear functions, IEEE Trans. Inf. Theory 55 (1) (2009) 70–82, http://dx.doi.org/10.1109/TIT.2008.2008145.
[23] A. Pott, Y. Zhou, Switching construction of planar functions on finite fields, in: M.A. Hasan, T. Helleseth (Eds.), WAIFI, in: Lect. Notes Comput. Sci., vol. 6087, Springer, 2010, pp. 135–150.
[24] C. Carlet, C. Ding, J. Yuan, Linear codes from perfect nonlinear mappings and their secret sharing schemes, IEEE Trans. Inf. Theory 51 (6) (2005) 2089–2102, http://dx.doi.org/10.1109/TIT.2005.847722.
[25] K. Nyberg, On the construction of highly non-linear permutations, in: EUROCRYPT'92, in: Lect. Notes Comput. Sci., vol. 658, Springer-Verlag, 1993, pp. 92–98.
[26] L. Budaghyan, C. Carlet, A. Pott, New classes of almost bent and almost perfect nonlinear polynomials, IEEE Trans. Inf. Theory 52 (3) (2006) 1141–1152.
[27] L. Budaghyan, C. Carlet, CCZ-equivalence of bent vectorial functions and related constructions, Des. Codes Cryptogr. 59 (1–3) (2011) 69–87.
[28] G. Kyureghyan, A. Pott, Some theorems on planar mappings, in: J. von zur Gathen, J. Imaña, c. Koç (Eds.), Arithmetic of Finite Fields, in: Lect. Notes Comput. Sci., vol. 5130, Springer, Berlin, Heidelberg, 2008, pp. 117–122.

[29] J. Daemen, V. Rijmen, AES Proposal, National Institute of Standards and Technology, Rijndael, 2000.
[30] S.W. Golomb, Theory of transformation groups of polynomials over $GF(2)$ with applications to linear shift register sequences, Inf. Sci. 1 (1968) 87–109.
[31] T. Kasami, The weight enumerators for several classes of subcodes of the second order binary Reed–Muller codes, Inf. Control 18 (1971) 369–394.
[32] H. Dobbertin, Almost perfect nonlinear power functions on $GF(2^n)$: the Niho case, Inf. Comput. 151 (1–2) (1999) 57–72.
[33] A. Canteaut, P. Charpin, H. Dobbertin, Binary m-sequences with three-valued crosscorrelation: a proof of Welch conjecture, IEEE Trans. Inf. Theory 46 (1) (2000) 4–8.
[34] H. Dobbertin, Almost perfect nonlinear power functions on $GF(2^n)$: the Welch case, IEEE Trans. Inf. Theory 45 (4) (1999) 1271–1275.
[35] H.D. Hollmann, Q. Xiang, A proof of the Welch and Niho conjectures on crosscorrelations of binary m-sequences, Finite Fields Appl. 7 (2) (2001) 253–286.
[36] T. Beth, C. Ding, On almost perfect nonlinear permutations, in: EUROCRYPT'93, in: Lect. Notes Comput. Sci., vol. 765, Springer, 1993, pp. 65–76.
[37] H. Dobbertin, Almost perfect nonlinear power functions on $GF(2^n)$: a new class for n divisible by 5, in: Proceedings of Finite Fields and Applications Fq5, Springer-Verlag, Augsburg, Germany, 2000, pp. 113–121.
[38] R.S. Coulter, Rex W. Matthews, Planar functions and planes of Lenz–Barlotti class, II, Des. Codes Cryptogr. 10 (1997) 167–184.
[39] F. Hernando, G. McGuire, Proof of a conjecture on the sequence of exceptional numbers, classifying cyclic codes and APN functions, J. Algebra 343 (1) (2011) 78–92.
[40] F. Hernando, G. McGuire, F. Monserrat, On the classification of exceptional planar functions over $\mathbb{F}_p$, CoRR abs/1301.4016, http://arxiv.org/abs/1301.4016, 2013.
[41] E. Leducq, Functions which are PN on infinitely many extensions of $\mathbb{F}_p$, p odd, Des. Codes Cryptogr. (2014), http://dx.doi.org/10.1007/s10623-013-9912-6, in press, http://arxiv.org/abs/1006.2610.
[42] M.E. Zieve, Planar functions and perfect nonlinear monomials over finite fields, Des. Codes Cryptogr., in press, http://arxiv.org/pdf/1301.5004.pdf.
[43] F. Caullery, K.-U. Schmidt, Y. Zhou, Exceptional planar polynomials, Des. Codes Cryptogr. (2014), in press, http://arxiv.org/abs/1403.4015.
[44] Y. Edel, G. Kyureghyan, A. Pott, A new APN function which is not equivalent to a power mapping, IEEE Trans. Inf. Theory 52 (2) (2006) 744–747, http://dx.doi.org/10.1109/TIT.2005.862128.
[45] L. Budaghyan, C. Carlet, G. Leander, Constructing new APN functions from known ones, Finite Fields Appl. 15 (2) (2009) 150–159.
[46] K. Nyberg, S-boxes and round functions with controllable linearity and differential uniformity, in: B. Preneel (Ed.), FSE, in: Lect. Notes Comput. Sci., vol. 1008, Springer, 1994, pp. 111–130.
[47] X.-d. Hou, Affinity of permutations of $\mathcal{F}_2^n$, Discrete Appl. Math. 154 (2) (2006) 313–325.
[48] K.A. Browning, J.F. Dillon, R. Kibler, M. McQuistan, APN polynomials and related codes, J. Comb. Inf. Syst. Sci. 34 (2010), Amer. Math. Soc..
[49] Y. Edel, A. Pott, A new almost perfect nonlinear function which is not quadratic, Adv. Math. Commun. 3 (1) (2009) 59–81.
[50] Y. Yu, M. Wang, Y. Li, A matrix approach for constructing quadratic APN functions, in: M.G.P. Lilya Budaghyan, Tor Helleseth (Eds.), Pre-Proceedings of WCC 2013, 2013, pp. 48–57.
[51] R.S. Coulter, F. Lazebnik, On the classification of planar monomials over fields of square order, Finite Fields Appl. 18 (2) (2012) 316–336.
[52] R.M. Hakala, Results on linear models in cryptography, PhD thesis, Aalto University School of Science, 2013, http://urn.fi/URN:ISBN:978-952-60-5027-0.
[53] X. Lai, J.L. Massey, A proposal for a new block encryption standard, in: I. Damgård (Ed.), EUROCRYPT'90, in: Lect. Notes Comput. Sci., vol. 473, Springer, 1991, pp. 389–404.
[54] N. Brandstätter, T. Lange, A. Winterhof, On the non-linearity and sparsity of Boolean functions related to the discrete logarithm in finite fields of characteristic two, in: Ø. Ytrehus (Ed.), WCC 2005, in: Lect. Notes Comput. Sci., vol. 3969, Springer, 2006, pp. 135–143.
[55] C. Carlet, K. Feng, An infinite class of balanced vectorial Boolean functions with optimum algebraic immunity and good nonlinearity, in: Y.M. Chee, C. Li, S. Ling, H. Wang, C. Xing (Eds.), IWCC 2009, in: Lect. Notes Comput. Sci., vol. 5557, Springer, 2009, pp. 1–11.
[56] R.M. Hakala, K. Nyberg, On the nonlinearity of discrete logarithm in $\mathbb{F}_{2^n}$, in: C. Carlet, A. Pott (Eds.), SETA 2010, in: Lect. Notes Comput. Sci., vol. 6338, Springer, 2010, pp. 333–345.

[57] D. Tang, C. Carlet, X. Tang, Highly nonlinear boolean functions with optimal algebraic immunity and good behavior against fast algebraic attacks, IEEE Trans. Inf. Theory 59 (1) (2013) 653–664.
[58] Q. Wang, C.H. Tan, Properties of a family of cryptographic boolean functions, in: SETA 2014, in press.
[59] K. Drakakis, R. Gow, G. McGuire, APN permutations on $\mathbb{Z}_n$ and Costas arrays, Discrete Appl. Math. 157 (15) (2009) 3320–3326.
[60] J.L. Massey, SAFER K-64: a byte-oriented block-ciphering algorithm, in: R.J. Anderson (Ed.), Fast Software Encryption, FSE'93, in: Lect. Notes Comput. Sci., vol. 809, Springer, 1994, pp. 1–17.
[61] J.L. Massey, SAFER K-64: one year later, in: B. Preneel (Ed.), Fast Software Encryption, FSE'94, in: Lect. Notes Comput. Sci., vol. 1008, 1995, pp. 212–241.
[62] K. Drakakis, V. Requena, G. McGuire, On the nonlinearity of exponential Welch Costas functions, IEEE Trans. Inf. Theory 56 (3) (2010) 1230–1238.
[63] R.M. Hakala, An upper bound for the linearity of Exponential Welch Costas functions, Finite Fields Appl. 18 (4) (2012) 855–862.
[64] G. Piret, T. Roche, C. Carlet, PICARO - A block cipher allowing efficient higher-order side-channel resistance, in: F. Bao, P. Samarati, J. Zhou (Eds.), ACNS, in: Lect. Notes Comput. Sci., vol. 7341, Springer, 2012, pp. 311–328.
[65] P. Charpin, G.M. Kyureghyan, A note on verifying the APN property, Cryptology ePrint Archive, Report 2013/475 http://eprint.iacr.org/, 2013.
[66] K. Nyberg, L.R. Knudsen, Provable security against differential cryptanalysis, in: E.F. Brickell (Ed.), CRYPTO, in: Lect. Notes Comput. Sci., vol. 740, Springer, 1992, pp. 566–574.
[67] K. Nyberg, L.R. Knudsen, Provable security against a differential attack, J. Cryptol. 8 (1) (1995) 27–37.
[68] K. Aoki, On maximum non-averaged differential probability, in: S.E. Tavares, H. Meijer (Eds.), Selected Areas in Cryptography, in: Lect. Notes Comput. Sci., vol. 1556, Springer, 1998, pp. 118–130.
[69] T. Jakobsen, L.R. Knudsen, The interpolation attack on block ciphers, in: E. Biham (Ed.), FSE, in: Lect. Notes Comput. Sci., vol. 1267, Springer, 1997, pp. 28–40.
[70] C. Boura, A. Canteaut, On the influence of the algebraic degree of F^{-1} on the algebraic degree of $G \circ F$, IEEE Trans. Inf. Theory 59 (1) (2013) 691–702.
[71] M. Matsui, New structure of block ciphers with provable security against differential and linear cryptanalysis, in: D. Gollmann (Ed.), FSE, in: Lect. Notes Comput. Sci., vol. 1039, Springer, 1996, pp. 205–218.
[72] M. Matsui, New block encryption algorithm MISTY, in: E. Biham (Ed.), FSE, in: Lect. Notes Comput. Sci., vol. 1267, Springer, 1997, pp. 54–68.
[73] V. Niemi, K. Nyberg, UMTS Security, Wiley & Sons, 2003.
[74] J. Daemen, L.R. Knudsen, V. Rijmen, Linear frameworks for block ciphers, Des. Codes Cryptogr. 22 (1) (2001) 65–87, http://dx.doi.org/10.1023/A:1008303310011.
[75] J. Daemen, V. Rijmen, Understanding two-round differentials in AES, in: R.D. Prisco, M. Yung (Eds.), SCN, in: Lect. Notes Comput. Sci., vol. 4116, Springer, 2006, pp. 78–94.
[76] J. Daemen, V. Rijmen, Plateau characteristics, IET Inform. Secur. 1 (1) (2007) 11–17.
[77] H. Gilbert, T. Peyrin, Super-Sbox cryptanalysis: improved attacks for AES-like permutations, in: S. Hong, T. Iwata (Eds.), FSE, in: Lect. Notes Comput. Sci., vol. 6147, Springer, 2010, pp. 365–383.
[78] A. Canteaut, J. Roué, Extended differential properties of cryptographic functions, in: The 11th International Conference on Finite Fields and Their Applications, in: Contemp. Math., Amer. Math. Soc., Magdeburg, Germany, July 2013.
[79] G. Leander, A. Poschmann, On the classification of 4 bit S-boxes, in: C. Carlet, B. Sunar (Eds.), WAIFI, in: Lect. Notes Comput. Sci., vol. 4547, Springer, 2007, pp. 159–176.
[80] M. Brinkmann, G. Leander, On the classification of APN functions up to dimension five, Des. Codes Cryptogr. 49 (1–3) (2008) 273–288.
[81] J. Borghoff, A. Canteaut, T. Güneysu, E.B. Kavun, M. Knezevic, L.R. Knudsen, G. Leander, V. Nikov, C. Paar, C. Rechberger, P. Rombouts, S.S. Thomsen, T. Yalçin, PRINCE – a low-latency block cipher for pervasive computing applications – extended abstract, in: X. Wang, K. Sako (Eds.), ASIACRYPT, in: Lect. Notes Comput. Sci., vol. 7658, Springer, 2012, pp. 208–225.
[82] B. Gérard, V. Grosso, M. Naya-Plasencia, F.-X. Standaert, Block ciphers that are easier to mask: how far can we go?, in: G. Bertoni, J.-S. Coron (Eds.), CHES, in: Lect. Notes Comput. Sci., vol. 8086, Springer, 2013, pp. 383–399.
[83] B. Bilgin, A. Bogdanov, M. Knezevic, F. Mendel, Q. Wang, Fides: Lightweight authenticated cipher with side-channel resistance for constrained hardware, in: G. Bertoni, J.-S. Coron (Eds.), CHES, in: Lect. Notes Comput. Sci., vol. 8086, Springer, 2013, pp. 142–158.

[84] M. Ullrich, C. De Cannière, S. Indesteege, O. Küçük, N. Mouha, B. Preneel, Finding optimal bitsliced implementations of 4 × 4-bit S-boxes, in: G. Leander, S.S. Thomsen (Eds.), SKEW 2011, 2011.
[85] K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima, T. Tokita, Camellia: a 128-bit block cipher suitable for multiple platforms – design and analysis, in: D.R. Stinson, S.E. Tavares (Eds.), Selected Areas in Cryptography, in: Lect. Notes Comput. Sci., vol. 2012, Springer, 2000, pp. 39–56.
[86] A. Bogdanov, L.R. Knudsen, G. Leander, C. Paar, A. Poschmann, M.J.B. Robshaw, Y. Seurin, C. Vikkelsoe, PRESENT: an ultra-lightweight block cipher, in: P. Paillier, I. Verbauwhede (Eds.), CHES, in: Lect. Notes Comput. Sci., vol. 4727, Springer, 2007, pp. 450–466.
[87] J. Guo, T. Peyrin, A. Poschmann, M.J.B. Robshaw, The LED block cipher, in: B. Preneel, T. Takagi (Eds.), CHES'11, in: Lect. Notes Comput. Sci., vol. 6917, Springer, 2011, pp. 326–341.
[88] C. Blondeau, A. Canteaut, P. Charpin, Differential properties of $x \mapsto x^{2^t-1}$, IEEE Trans. Inf. Theory 57 (12) (2011) 8127–8137.
[89] C. Blondeau, A. Canteaut, P. Charpin, Differential properties of power functions, Int. J. Inform. Coding Theory 1 (2) (2010) 149–170.
[90] C. Blondeau, L. Perrin, More differentially 6-uniform power functions, Des. Codes Cryptogr. 73 (2) (2014) 487–505.
[91] K. Nyberg, Generalized Feistel networks, in: K. Kim, T. Matsumoto (Eds.), ASIACRYPT, in: Lect. Notes Comput. Sci., vol. 1163, Springer, 1996, pp. 91–104.
[92] N. Courtois, J. Pieprzyk, Cryptanalysis of block ciphers with overdefined systems of equations, in: Y. Zheng (Ed.), ASIACRYPT'02, in: Lect. Notes Comput. Sci., vol. 2501, Springer, 2002, pp. 267–287.
[93] C. Cid, G. Leurent, An analysis of the XSL algorithm, in: B.K. Roy (Ed.), ASIACRYPT'05, in: Lect. Notes Comput. Sci., vol. 3788, Springer, 2005, pp. 333–352.
[94] G. Bertoni, J. Daemen, M. Peeters, G.V. Assch, Keccak sponge function family main document, submission to NIST (Round 3), 2011.
[95] I. Dinur, J. Jean, Cryptanalysis of FIDES, in: C. Cid, C. Rechberger (Eds.), Fast Software Encryption, FSE 2014, Springer-Verlag, in press.
[96] F.-X. Standaert, G. Piret, G. Rouvroy, J.-J. Quisquater, J.-D. Legat, Iceberg: an involutional cipher efficient for block encryption in reconfigurable hardware, in: B.K. Roy, W. Meier (Eds.), FSE, in: Lect. Notes Comput. Sci., vol. 3017, Springer, 2004, pp. 279–299.
[97] C. Carlet, Relating three nonlinearity parameters of vectorial functions and building APN functions from bent functions, Des. Codes Cryptogr. 59 (1–3) (2011) 89–109.
[98] A. Canteaut, M. Naya-Plasencia, Structural weaknesses of permutation with a low differential uniformity and generalised croked functions, in: Contemp. Math., vol. 518, Amer. Math. Soc., 2010, pp. 55–71.
[99] T. Bending, D. Fon-Der-Flaass, Crooked functions, bent functions, and distance regular graphs, Electron. J. Comb. 5 (1) (1998).
[100] R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks, L. Wingers, The SIMON and SPECK families of lightweight block ciphers, Cryptology ePrint Archive, Report 2013/404, http://eprint.iacr.org/, 2013.